Discussion of “On simulation and properties of the stable law” by L. Devroye and L. James

Published version:

DOI:10.1007/s10260-014-0269-4

Statistical Methods and Applications manuscript No.
(will be inserted by the editor)

Discussion of "On simulation and properties of the stable law" by L. Devroye and L. James

Antonio Lijoi · **Igor Prünster**

Received: date / Accepted: date

We wish to congratulate the authors for their brilliant and remarkable piece of work. This is an excellent overview of methods for generating random variables associated to the stable distribution. We are also impressed by the number of distributional identities they have singled out and by the useful and systematic account of the literature in the field. In addition to being of great interest to Statistics and Probability, the contents of the paper actually are valuable to a number of research areas beyond these.

As pointed out by the authors, distributional results, and the associated simulation schemes, for S_α, with $\alpha \in (0,1)$, $L_{\rho,p}$ and M_α are of great usefulness in statistical applications. In particular, a number of results displayed in the paper are relevant for Bayesian nonparametric inference and are closely connected with a considerable portion of research we have been working on in the past few years. Here we will focus on two main issues: (a) the use of some simulation strategies to generate random probabilities, based on α–stable subordinators or, more generally, completely random measures; (b) the determination of the probability distribution of linear functionals of Poisson–Dirichlet random measures. Another noteworthy issue that will not be touched upon concerns species sampling models where polynomially and exponentially tilted stable random variates and Mittag–Leffler distributions arise as a result

A. Lijoi and I. Prünster are also affiliated to Collegio Carlo Alberto, Moncalieri, Italy. Support from the European Research Council (ERC) through StG "N-BNP" 306406 is gratefully acknowledged.

Antonio Lijoi
Department of Economics and Management
University of Pavia
E-mail: lijoi@unipv.it

Igor Prünster
Department of Economics and Statistics
University of Torino
E-mail: igor.pruenster@unito.it

of limiting procedures on the number of distinct species as the sample size diverges. This topic will be the focus of the allied contribution by S. Favaro and B. Nipoti to the discussion and completes the picture on the role unilateral α–stable distributions have in nowadays Bayesian nonparametric inference.

In the sequel, for any measure m on a separable and complete metric space we set $m(f) := \int_{\mathbb{X}} f(x)\, m(\mathrm{d}x)$, where $f : \mathbb{X} \to \mathbb{R}^+$ is any measurable function such that $m(f) < \infty$ (almost surely).

1 Random probabilities based on α–stable laws

In Bayesian nonparametric inference, random probability measures represent a fundamental tool, given their laws can be interpreted as prior distributions on spaces of probability measures. The most celebrated instance is the Dirichlet process, which stands out thanks to its analytical tractability. See [9]. Some relevant generalizations of the Dirichlet process that have appeared in the literature arise in connection with unilateral α–stable distributions. Examples include the normalized α–stable random measure, the two parameter Poisson–Dirichlet process, the normalized generalized gamma process and Gibbs–type priors. Recent reviews are provided in [21, 6].

1.1 α–stable random measure and related priors

In fact, the definition of the four above mentioned classes of nonparametric priors can be based on the use of α–stable completely random measures (CRMs), which we denote by $\tilde{\mu}_\alpha$. Since we use $\mathbb{1}_A$ to denote the indicator function of set A, $\tilde{\mu}_\alpha(\mathbb{1}_A)$ is the random mass $\tilde{\mu}_\alpha$ assigns to set A. Hence, $\tilde{\mu}_\alpha$ is a random element taking values in the space $M_{\mathbb{X}}$ of boundedly finite measures on a separable and complete metric space $\mathbb{X}$, such that the random variables $\tilde{\mu}_\alpha(\mathbb{1}_{A_1}), \ldots, \tilde{\mu}_\alpha(\mathbb{1}_{A_d})$ are independent for any choice of pairwise disjoint measurable sets $A_1, \ldots, A_d$ and

$$\mathbb{E}\left[\exp\left\{-\tilde{\mu}_\alpha(f)\right\}\right] = \exp\left\{-G(f^\alpha)\right\} \tag{1}$$

where G is a σ–finite measure on $\mathbb{X}$. Hence, an α–stable CRM, $\tilde{\mu}_\alpha$, is characterized by its parameter α and the measure G. Note that, if G is finite then $\tilde{\mu}_\alpha(\mathbb{1}_{\mathbb{X}}) < \infty$ (a.s.). In the sequel G will be assumed to be a non–atomic probability measure. Within this setup, one can easily define the corresponding random probability measure via normalization as $\tilde{p}_{\alpha,0} := \tilde{\mu}_\alpha / \tilde{\mu}_\alpha(\mathbb{1}_{\mathbb{X}})$. Simulation of the trajectories of $\tilde{p}_{\alpha,0}$ is typically achieved by relying on its stick–breaking representation [22] or, alternatively, by means of the Ferguson–Klass series representation [10] according to which

$$\tilde{\mu}_\alpha = \sum_{i \geq 1} J_i\, \delta_{X_i} \tag{2}$$

with the X_i's being iid from G and independent from the ordered jumps $J_1 > J_2 > \cdots$. Furthermore the jumps can be simulatedho visto che qui hai fatto una frase a parte, ma secondo me non va bene perche' per parlare di rappresentazione di FK e' necessario dire che i salti sono ottenuti mediante inversione delle intensita' (e' quella la FK!). se vuoi possiamo tenere in frasi diverse la simulazione, ma qualcosa nella precedente bisogna dire sul legame tra inversione intensita' e J_i's by inverting the Lévy intensity associated to $\tilde{\mu}_\alpha$. On the other hand, if interest lies in evaluating $\tilde{p}_{\alpha,0}$ on some specific set, i.e. $\tilde{p}_{\alpha,0}(\mathbb{1}_A)$, the paper provides readily usable simulation tools. In fact, (1) trivially implies that $\tilde{\mu}_\alpha(\mathbb{1}_A) \stackrel{\mathcal{L}}{=} G(\mathbb{1}_A)^{1/\alpha}\, S_\alpha$ and one can resort to random variate generators for S_α in order to simulate $\tilde{\mu}_\alpha(\mathbb{1}_A)$, for any set A. Thus,

$$\tilde{p}_{\alpha,0}(\mathbb{1}_A) = \frac{\tilde{\mu}_\alpha(\mathbb{1}_A)}{\tilde{\mu}_\alpha(\mathbb{1}_A) + \tilde{\mu}_\alpha(\mathbb{1}_{A^c})}$$

so that it is enough to simulate independent stable random variates $\tilde{\mu}_\alpha(\mathbb{1}_A)$ and $\tilde{\mu}_\alpha(\mathbb{1}_{A^c})$. This also carries over to the analysis of the posterior distribution of $\tilde{p}_\alpha$. Indeed, if $(X_n)_{n\ge 1}$ is a sequence of exchangeable random elements taking values in $\mathbb{X}$ such that $(X_i \,|\, \tilde{p}_\alpha) \stackrel{\text{iid}}{\sim} \tilde{p}_\alpha$, a representation of the posterior distribution of $\tilde{\mu}_\alpha$ in terms of mixtures of exponentially tilted α–stable CRMs can be given. See [16]. To be more precise, if U_n is a random variable such that U_n^α is gamma distributed with shape and scale parameters equal to k and 1, respectively, then the posterior distribution of $\tilde{\mu}_\alpha$ given $U_n = u$ and the observations $X_1, \ldots, X_n$ equals the probability distribution of the random measure

$$\tilde{\mu}_{\alpha,u} + \sum_{i=1}^{k} J_{i,u}\, \delta_{x_j^*} \tag{3}$$

Here, $x_1^*, \ldots, x_k^*$ are the k distinct values in the sample $X_1, \ldots, X_n$ with respective frequencies $n_1, \ldots, n_k$, the jumps $J_{i,u}$'s are independent and gamma distributed with shape parameter $n_i - \alpha$ and scale parameter u and $\tilde{\mu}_{\alpha,u}$ is a CRM such that

$$\mathbb{E}\left[\exp\left\{-\tilde{\mu}_{\alpha,u}(f)\right\}\right] = \exp\left\{-G((f + u\mathbb{1}_{\mathbb{X}})^\alpha) + u^\alpha\right\}$$

and is also known as generalized gamma process. See [1]. Moreover, the jumps $(J_{i,u})_i$ and $\tilde{\mu}_{\alpha,u}$ are independent. In this case, the simulation of $\tilde{p}_{\alpha,u}(\mathbb{1}_A) = \tilde{\mu}_{\alpha,u}(\mathbb{1}_A)/\tilde{\mu}_{\alpha,u}(\mathbb{1}_{\mathbb{X}})$ amounts to simulating independent exponentially tilted unilateral α–stable variables, which are referred to as $(S_\alpha)_u^*$ in the paper. An efficient rejection method for their simulation is devised in [7].

It should be noted that $\tilde{\mu}_\alpha$ is also the key ingredient for the definition of the Poisson–Dirichlet process with parameters $(\alpha, \theta) \in [0, 1) \times (-\alpha, \infty)$ that has been recently popularized as the Pitman–Yor process. In order to define such a process, let $\tilde{\mu}_{\alpha,\theta}$ be a random measure on $\mathbb{X}$ whose probability distribution on $M_{\mathbb{X}}$, say $\mathbb{P}_{\alpha,\theta}$, is absolutely continuous with respect to the probability distribution $\mathbb{P}_\alpha$ of $\tilde{\mu}_\alpha$, then $(\mathrm{d}\mathbb{P}_{\alpha,\theta}/\mathrm{d}\mathbb{P}_\alpha)(m) = m(\mathbb{1}_{\mathbb{X}})^{-\theta}\, \Gamma(\theta+1)/\Gamma(\theta/\sigma+1)$. A Pitman–Yor process with parameters (α, θ), then, corresponds to a random probability measure obtained by normalizing $\tilde{\mu}_{\alpha,\theta}$, i.e. $\tilde{p}_{\alpha,\theta} = \tilde{\mu}_{\alpha,\theta}/\tilde{\mu}_{\alpha,\theta}(\mathbb{1}_{\mathbb{X}})$. The normalized α–stable random measure is, then, a special case corresponding to $(\alpha, \theta) = (\alpha, 0)$. As shown in [21], a representation of type (3) still holds true and, accordingly, one can simulate exponentially tilted unilateral α–stable variables for evaluating a posterior Pitman–Yor process on any given set. Pitman-Yor processes are nowadays highly popular in Bayesian nonparametrics practice thanks to the development of efficient simulation algorithms that allow to sample from the posterior in complex hierarchical mixture models. See, e.g., [13, 24, 16]. ho tolto il nostro (citato poche righe prima) e aggiunto Stephen: però quello di Stephen si riferisce al Dirichlet anche se sappiamo può essere esteso a stabile e affini. Lo diciamo questo, visto che il titolo del lavoro stesso enfatizza il Dirichlet? ho aggiunto anche il nostro con lancillotto cosi' stempera la questione che quello di steve enfatizzi l'MDP e dunque lascerei cosi'.

1.2 Multivariate α–stable random measures

As the authors concisely mention, a relevant "open problem" is given by the investigation of multivariate stable laws and related simulation algorithms. In connection to the above discussion, this may be rephrased as the problem of specifying vectors of dependent α–stable CRMs that are still analytically tractable to the extent of allowing the evaluation of Bayesian inferences, at least approximately via suitable sampling schemes. These have actually become of interest in Bayesian inference since they are useful to define models that are able to cope with forms of dependence more general than exchangeability. A recent proposal in [18, 19] induces dependence by considering marginal α–stable CRMs $\tilde{\mu}_{i,\alpha} = \tilde{\mu}_i + \tilde{\mu}_0$, for $i = 1, 2$, where $\tilde{\mu}_1$, $\tilde{\mu}_2$ and $\tilde{\mu}_0$ are independent α–stable CRMs characterized by the following Laplace functional transforms

$$\begin{aligned} \mathbb{E}\left[\exp\left\{-\tilde{\mu}_i(f)\right\}\right] &= \exp\left\{-z\, G(f^\alpha)\right\} \qquad (i = 1, 2), \\ \mathbb{E}\left[\exp\left\{-\tilde{\mu}_0(f)\right\}\right] &= \exp\left\{-(1-z)\, G(f^\alpha)\right\} \end{aligned}$$

for any $z \in [0, 1]$. Each $\tilde{\mu}_{i,\alpha}$ has, then, Laplace functional transform coinciding with (1) and dependence is due to the sharing of a common component $\tilde{\mu}_0$. Such a construction is inspired by [11] where a characterization of pairs of *canonically correlated* Poisson random measures is provided. The main advantage of this construction is represented by the availability, under suitable assumptions, of a simple expression for the joint Laplace transform of the underlying correlated Poisson random measures.

The resulting normalized CRMs admit an intuitive mixture representation as

$$\tilde{p}_{i,\alpha} = \frac{\tilde{\mu}_{i,\alpha}}{\tilde{\mu}_{i,\alpha}(\mathbb{1}_{\mathbb{X}})} = w_i \frac{\tilde{\mu}_i}{\tilde{\mu}_i(\mathbb{1}_{\mathbb{X}})} + (1 - w_i) \frac{\tilde{\mu}_0}{\tilde{\mu}_0(\mathbb{1}_{\mathbb{X}})}$$

for $i = 1, 2$. One can envisage that random variate generators as those proposed in the paper may lead to quite a straightforward simulation of vectors of $(\tilde{p}_{1,\alpha}(\mathbb{1}_A), \tilde{p}_{2,\alpha}(\mathbb{1}_B))$. The only simulation techniques we are aware of in this setting are based on extensions of the representation in [10]. See [14] and [4], where the former refers only to the gamma case whereas the latter applies to the case where dependence is induced through Lévy copulas. Hence, there is no doubt that the implementation of dependent processes for Bayesian inference would greatly benefit from further progress on the simulation of multidimensional stable laws.

2 Random means

The specification of a prior on the space $P_{\mathbb{X}}$ of probability distributions on $\mathbb{X}$ induces a prior on finite–dimensional features of the distribution of the data. A noteworthy example is offered by the mean. In other words, if the prior coincides with the probability distribution of a random probability measure $\tilde{p}$, one might be interested in identifying the distribution of

$$\tilde{p}(f) = \int_{\mathbb{X}} f(x)\, \tilde{p}(\mathrm{d}x)$$

The study of the probability distribution of $\tilde{p}(f)$, when $\tilde{p}$ is a Dirichlet process with base measure $c\,G$, where $c \in (0,\infty)$ and G is a probability measure as in the previous section, has been initiated in the pioneering papers by D.M. Cifarelli and E. Regazzini. See [2, 3]. Their approach relies on the inversion of the generalized Stieltjes transform of $\tilde{p}(f)$ which is determined through the following fundamental identity, also known as Markov–Krein or Cifarelli–Regazzini identity,

$$\mathbb{E}\left[\frac{1}{(z+\tilde{p}(f))^c}\right] = \exp\left\{-c\int_{\mathbb{X}} \log[z+f(x)]\, G(\mathrm{d}x)\right\} \tag{4}$$

for any $z \in \mathbb{C}$ such that $\mathrm{Im}(z) \neq 0$ if $\mathrm{Re}(z) < 0$. It is interesting to note that, though the determination of (4) and the inversion of the left–hand–side has been originally motivated by statistical arguments, interest in the probability distribution of $\tilde{p}(f)$ has emerged in other research areas ranging from combinatorics to special function theory. See also [5, 15, 20]. Since the Dirichlet process is intimately related to a gamma structure, we do not linger further on it and rather focus on its generalizations as displayed in Section 1 wherein the stable structure comes into play.

In fact, the results of the discussion paper are of interest for means of random probability measures defined through the transformation of a α–stable CRM. Here Pitman–Yor processes play a prominent role. For instance, $\tilde{p}_{\alpha,\theta}(f)$ is related to the excursions of a Bessel bridge under suitable choices of f ad θ. If $\theta = \alpha$ and $G = p\delta_1 + (1-p)\delta_0$, then $\int x\,\tilde{p}_{\alpha,\theta}(\mathrm{d}x) \stackrel{\mathcal{L}}{=} L_{\alpha,p}$ with $L_{\alpha,p}$ denoting the Lamperti's second law as in the paper. Nice distributional identities yield simple simulation algorithms for sampling from the probability distribution of $\int x\,\tilde{p}_{\alpha,\theta}(\mathrm{d}x)$. If $\theta \neq \alpha$, one has

$$\mathbb{E}\left[\frac{1}{\{\lambda + \int x\,\tilde{p}_{\alpha,\theta}(\mathrm{d}x)\}^{\theta+1}}\right] = \frac{p(1+\lambda)^{\alpha-1} + (1-p)\lambda^{\alpha-1}}{\{p(1+\lambda)^{\alpha} + (1-p)\lambda^{\alpha}\}^{\frac{\theta}{\alpha}+1}}.$$

This leads to more complicated expressions for the density of $\int x\,\tilde{p}_{\alpha,\theta}(\mathrm{d}x)$ and it would be interesting to have suitable simulation algorithms applicable also to this case. Furthermore, it would be of great interest to obtain an algorithm that applies also to the case when G does not necessarily coincides with a mixture of point masses in 0 and in 1.

Another important issue concerns the direct evaluation of the posterior distribution of $\tilde{p}_{\alpha,\theta}(f)$. For the normalized α–stable case, i.e. $\tilde{p}_{\alpha,0}(f)$, one can have a glimpse of the analytical hurdles associated to the determination of the posterior distribution from [17]. The availability of a simulation algorithm for the posterior would be of great help. The algorithm in [12] and the double CFTP method in [8] represent effective solutions to the problem when $\tilde{p}$ is a Dirichlet process and rely on its conjugacy. In contrast, we are not aware of any algorithm that can be used for the case of Pitman–Yor means, even for specific choices of α and θ. Some progress might actually be based on the fact that, conditionally on the data $X_1,\dots,X_n$ featuring k distinct values $X_1^*,\dots,X_k^*$ with respective frequencies $n_1,\dots,n_k$, the posterior distribution of the mean $\int f\,\mathrm{d}\tilde{p}_{\alpha,\theta}$ equals the distribution of

$$\sum_{i=1}^{k} w_i\, f(X_i^*) + w_{k+1} \int_{\mathbb{X}} f(x)\,\tilde{p}_{\alpha,\theta+k\alpha}(\mathrm{d}x)$$

where $(w_1,\dots,w_k)$ is a k–variate Dirichlet vector with parameters $(n_1 - \sigma,\dots,n_k-\sigma,\theta+k\sigma)$ and is independent from $\tilde{p}_{\alpha,\theta+k\alpha}$. See [23].

In conclusion, we would like to once more congratulate the authors for a fine and inspiring paper.

References

1. Brix, A. (1999). Generalized gamma measures and shot-noise Cox processes. *Adv. Appl. Probab.* 31:929–953.
2. Cifarelli, D.M. and Regazzini, E. (1979). A general approach to Bayesian analysis of nonparametric problems. The associative mean values within the framework of the Dirichlet process. II. (Italian) *Riv. Mat. Sci. Econom. Soc.* 2:95–111.
3. Cifarelli, D.M. and Regazzini, E. (1990). Distribution functions of means of a Dirichlet process. *Ann. Statist.* 18:429-442.
4. Cont, R. and Tankov, P. (2004). *Financial modelling with jump processes.* Chapman & Hall/CRC, Boca Raton, FL.

5. Diaconis, P. and Kemperman, J. (1996). Some new tools on Dirichlet priors. In *Bayesian statistics 5* (J.M. Bernardo, J.O. Berger, A.P. Dawid and A.F.M. Smith, eds.) 97–106. Oxford Univ. Press.
6. De Blasi, P., Favaro, S., Lijoi, A., Mena, R., Prünster, I. and Ruggiero, M. (2014). Are Gibbs-type priors the most natural generalization of the Dirichlet process? *IEEE Trans. Pattern Anal. Mach. Intell.* doi: 10.1109/TPAMI.2013.217
7. Devroye, L. (2009). Random variate generation for exponentially and polynomially tilted stable distributions. *ACM Trans. Model. Comput. Simul.* 19:18.
8. Devroye, L. and James, L.F. (2011). The double CFTP method. *ACM Trans. Model. Comput. Simul.* 21:1–20.
9. Ferguson, T.S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* 1:209–230.
10. Ferguson, T.S. and Klass, M.J. (1972). A representation of independent increment processes without Gaussian components. *Ann. Math. Statist.* 43:1634–1643.
11. Griffiths, R.C. and Milne, R.K. (1978). A class of bivariate Poisson processes. *J. Mult. Anal.* 8:380–395.
12. Guglielmi, A., Holmes, C.C. and Walker, S.G. (2002). Perfect simulation involving functionals of a Dirichlet process. *J. Computat. Graph. Statist.* 11:306–310.
13. Ishwaran, H. and James, L.F. (2001). Gibbs sampling methods for stick-breaking priors. *J. Amer. Statist. Assoc.* 96:161–173.
14. Ishwaran, H. and Zarepour, M. (2009). Series representations for multivariate generalized gamma processes via a scale invariance principle. *Statist. Sinica* 19:1665–1682.
15. James, L.F. (2006). Functionals of Dirichlet processes, the Cifarelli-Regazzini identity and beta-gamma processes. *Ann. Statist.* 33:647–660.
16. James, L.F., Lijoi, A. and Prünster, I. (2009). Posterior analysis for normalized random measures with independent increments. *Scand. J. Stat.* 36:76–97.
17. James, L.F., Lijoi, A. and Prünster, I. (2010). On the posterior distribution of classes of random means. *Bernoulli* 16:155–180.
18. Lijoi, A., Nipoti, B. and Prünster, I. (2014). Bayesian inference with dependent normalized completely random measures. *Bernoulli* doi: 10.3150/13-BEJ521.
19. Lijoi, A., Nipoti, B. and Prünster, I. (2014). Dependent mixture models: clustering and borrowing information. *Comput. Stat. Data Anal.* **71**:417–433.
20. Lijoi, A. and Prünster, I. (2009). Distributional properties of means of random probability measures. *Statist. Surveys* 3:47-95.
21. Lijoi, A. and Prünster, I. (2010). Models beyond the Dirichlet process. In *Bayesian Nonparametrics* (Hjort, N.L., Holmes, C.C. Müller, P., Walker, S.G. Eds.), Cambridge University Press, Cambridge, pp. 80–136.
22. Perman, M., Pitman, J. and Yor, M. (1992). Size-biased sampling of Poisson point processes and excursions. *Probab. Theory Related Fields* 92:21–39.

23. Pitman, J. (1996). Some developments of the Blackwell-MacQueen urn scheme. In *Statistics, Probability and Game Theory* (T.S. Ferguson, L.S. Shapley and J.B. MacQueen, Eds.). IMS Lecture Notes Monogr. Ser. Vol. 30, Inst. Math. Statist., Hayward, pp. 245–267.
24. Walker, S.G. (2007). Sampling the Dirichlet mixture model with slices. *Comm. Statist. Simulation Comput.* 36:45–54.